

DOCUMENT RESUME

ED 247 466

CG 017 626

AUTHOR Krinsky, Richard; Clarke, Trent R.
TITLE Hope Springs Eternal: Prediction of Academic Success by College Students.
PUB DATE 26 Apr 84
NOTE 13p.; Paper presented at the Annual Convention of the Rocky Mountain Psychological Association (Las Vegas, NV, April 25-28, 1984).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *College Students; Feedback; *Grade Prediction; Higher Education; Individual Differences

ABSTRACT

Previous research has suggested that good students (those above the median) are better than poor students (below the median) at estimating their academic performance. To examine the relationship between pre-test predictions and post-test estimations of academic performance, 30 college students were asked to predict their scores on each of four tests in a psychology course. Following each test they were asked to estimate their performance, and make predictions about the performance of the class as a whole. Students also received feedback on prediction accuracy for the class. Before the third examination, they were given instructions for forming their predictions and estimations. Results showed that, while students generally overestimated their test scores, better students were closer to the target. When explicitly forewarned about overestimation errors, students were successful in lowering their estimations for the next test; when not forewarned, they were unsuccessful in lowering their estimated scores. Students above and below the median predicted class average with the same degree of precision. Results seem to be consistent with indications that good and poor students form differential assessments of their own test performance, but are similar in their assessments of the performance of others. (JAC)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED247466

HOPE SPRINGS ETERNAL: PREDICTION OF ACADEMIC SUCCESS BY COLLEGE STUDENTS

RICHARD KRINSKY

TRENT R. CLARKE

UNIVERSITY OF SOUTHERN COLORADO

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

* This document has been reproduced as
received from the person or organization
originating it.
☐ Minor changes have been made to improve
reproduction quality.

• Points of view or opinions stated in this docu-
ment do not necessarily represent official NIE
position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Richard Krinsky

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

Paper presented to the Rocky Mountain Psychological Association
meeting, Las Vegas, Nevada, April, 26, 1984.

CG 017626

INTRODUCTION

In an investigation of the relationship between academic achievement and the self-evaluations of academic performance, Moreland, Miller and Laucka (1981) asked 52 students enrolled in an introductory course at the University of Michigan to estimate the number of points they would receive on an essay question that appeared on the midterm examination, and another essay question that appeared on the final examination. In addition to estimating the number of points that they would receive for the essay questions, subjects were also asked to recall the instructor's grading criteria in the correct weighted order (i.e., the first criteria being the content of the response, the second criteria being the creativity of the response and the third criteria being the organization of the response). On the day after each exam, students were given the (anonymous) essay responses of three classmates and were asked to estimate the number of points that each would receive from the instructor.

The results of this investigation were in-line with previous studies indicating that "good" students were better at estimating their academic performance than "poor" students (e.g., Biggs & Tinsley, 1970; Gadzella, et al., 1976; Keefer, 1971; Yates & Karmos, 1977). In addition, Moreland, Miller and Laucka (1981) discovered that while "poor" students (i.e., students scoring below the median on the midterm examination) were, in fact, less accurate than "good" students (i.e., students scoring above the median) in evaluating their respective course performance, they did not differ in their accuracy of evaluating the course performance of others. It appeared as though the "poor" students were equally familiar with the instructor's grading criteria and were successful in applying these criteria to the work of others, but failed to apply these same criteria to their own work.

One reason that "poor" students may have so much difficulty in evaluating their own academic performance is because of attributional variables. It may be the case, as suggested by Moreland, Miller and Laucka, (1981) that "poor" students are more likely than "good" students to attribute their academic performance to external factors such as task difficulty or luck than to internal factors such as ability or effort. If "poor" students believe that their own performance on examinations is determined primarily by external factors, then their estimation inaccuracy can be explained by their unwillingness to apply the instructor's criteria (which is related more to internal factors than to external factors) to their own efforts, but a willingness to apply these criteria to the efforts of others.

Another reason that "poor" students may have had difficulty in evaluating their own performance but not the performance of others is because of test anxiety. It is possible, according to Moreland, Miller and Laucka (1981) that "poor" students suffered from an elevated level of anxiety immediately after the examination and were therefore handicapped in forming an accurate assessment of their essay performance. On the following day, under a more relaxed atmosphere, these students were able to accurately assess the essay performance of others.

While these post-hoc explanations of Moreland, Miller and Laucka (1981) are certainly plausible, other, more parsimonious explanations might be formulated as a-priori hypotheses for future experiments. For example, perhaps one's predictions for success (prior to an examination) might influence their estimation of performance (after an examination). It may be the case that "poor" students have special problems assessing their academic performance because they form unrealistic predictions about the eventual likelihood for their success. That issue was examined by Morrison and Morrison in 1979. They reported that students with high self-esteem relative to students with low self-esteem predicted higher grades before an examination, however, there were no reliable differences in their estimations following an examination. Interestingly, they also reported, contrary to Moreland, Miller and Laucka (1981) that students with external locus of control were more accurate than internals in estimating their grades. One problem with the Morrison and Morrison (1979) study was that subjects either predicted their scores before an examination or following an examination, therefore, only between-subjects, not within-subjects (prediction versus estimation) effects were reported.

Another reason that "poor" students may have special problems in estimating their academic performance is because of the cognitive complexity of the task. It may simply take those students performing below the median more trials to learn to estimate their performance with the same level of accuracy as the students performing above the median. There is evidence for this assertion. Moreland, Miller and Laucka (1981) reported that, "Overall, students become more accurate in their performance evaluations as the semester progressed" (pp. 339). Since, in the Moreland, Miller and Laucka (1981) study, there were only two essay questions evaluated (one from the midterm and another from the final), any differences due to practice effects, for students performing above and below the median, would require additional examinations to detect.

It is our purpose, in the following experiment, to examine the relationships

between the pre-test predictions and post-test estimations of academic performance for undergraduate students enrolled in a typical undergraduate course. We were interested in examining both self-assessment of academic performance, as well as, the academic assessment for others. In addition, we tested for individual differences for students performing above and below the overall test-score median and for differences due to practice effects.

EXPERIMENT #1

Students enrolled in a general psychology course (Summer Semester 1983, N = 30) served as subjects for this experiment. During the semester four (4) examinations were given (an examination was given after every two-week period) and served as the basis for the course grades. All examinations were similar in that they were equally weighted (i.e., worth 60 points each) and composed of objective and short- and long-essay type subsections. The bulk of each examination was objective in nature. Approximately 40% of each examination was formed by standard 4-alternative multiple choice questions and approximately 25% was formed by matching or true/false type questions. In addition, each examination had a section describing a psychological experiment (not found in the text) that required the students to identify independent and dependent variables, experimental designs and prepare graphs for all significant and nonsignificant effects. These sections accounted for approximately 25% of each examination. The remaining 10% of each examination included either labeling, definitional, fill-in or short-essay type questions. Overall, approximately 65% of every examination was objective in nature.

The experimental procedure was quite simple. Before every examination students were asked to form predictions about their anticipated test scores and following every test they were asked to form estimations about their performance. After the first examination students were asked to form pre-test predictions and post-test estimations of the class average and they were also provided with an incentive for the post-test accuracy of their estimated examination scores. The incentive was either a 2-point bonus (for estimating within plus or minus 2 points), or, a 1-point bonus (for estimating within plus or minus 4-points of their actual score). Before the third examination all students were given the feedback regarding the general level of estimation accuracy for the class, and given specific instructions for forming their test 3 predictions and estimations. In other words, all students were given their personal test-score feedback along with the distribution of all other scores in the

class after the first examination and they were given feedback regarding their estimation accuracy after the second examination. Students formed their test 4 predictions and estimations under conditions similar to test 2. They were given feedback from their previous examinations and an incentive for forming accurate post-test estimations for their actual test 4 scores. They were not, however, given specific instructions for forming their test 4 estimations.

At the first level of analysis, a separate repeated measures analysis of variance (ANOVA), across all examinations, for pre-test predicted scores (PREDESC), actual obtained scores (ACTSC) and post-test estimated scores (ESTSC) was conducted. The average scores for each of those dependent measures of performance, the F values and the results of all pairwise comparisons are summarized in the following table.*

TABLE 1
PREDICTIONS AND ESTIMATIONS OF GENERAL PSYCHOLOGY EXAMINATION
SCORES
SUMMER 1983, N=30

RESPONSES	TEST 1	TEST 2	TEST 3	TEST 4	F	COMPARISONS
PREDSC	49.19	46.77	44.38	44.48	P<.001	1>2>3=4
ACTSC	44.95	41.66	43.51	39.70	P<.001	1=3>2=4
ESTSC	47.45	46.51	42.46	45.16	P<.005	3<1=2=4

*All pairwise comparisons were derived by the Duncans Multiple Range Test (Edwards, p. 130). All differences were $p<.05$

Our results, at this level of analysis suggested that there were reliable differences within each of these performance measures and that the pattern of differences varied from one dependent measure to another. For example, the pattern for predicted scores indicated that students predicted reliably higher scores for test 1 than for test 2 and they predicted reliably higher scores for test 2 than for test 3 (tests 3 and 4 were not reliably different from each other).

The pattern of obtained test scores indicated that the level of difficulty, across examinations, was not homogenous. Reliable differences between average examination scores was detected. Tests 1 and 3 did not differ reliably from each other but were reliably greater than tests 2 and 4. Therefore, it appeared that tests 1 and 3, whose combined average of 44.23 (74%) were less rigorous than tests 2 and 4, whose combined average was 40.68 (68%).

The pattern for post-test estimations produced only one reliable difference. The average estimated score for test 3 (one of the easier examinations) was reliably

lower than any of the other tests and was the only correct estimation for any of the examination averages.

Wilcoxon Signed Rank tests (Mosteller & Rourke, p. 89) were conducted for differences among the various performance measures for each examination. These results are presented in table 2.

TABLE 2
WILCOXON SIGNED RANK TESTS BETWEEN PREDICTED, ACTUAL AND ESTIMATED
PERFORMANCE SCORES FOR GENERAL PSYCHOLOGY EXAMINATIONS
SUMMER 1983, N=30

	PKEDSC VS ACTSC	PREDSC VS ESTSC	ESTSC VS ACTSC
Test 1	308*	138	209*
Test 2	333*	59	352*
Test 3	34	260*	92
Test 4	349*	12	383*

* $P < .05$, $D_{crit.} = 189$

Our results were straightforward in that both the pre-test predictions and post-test estimations varied reliably from the actual performance measure for every examination, with the exception of test 3. Interestingly, while those performance estimates were reliably greater than the actual level of performance, they did not vary reliably from each other. This was true for tests 1, 2 and 4. It was only for test 3 that students were accurate in predicting their actual test score performance. In this case, they were accurate with both their pre-test predictions and their post-test estimations. The conditions for test 3 were such that students predicting and estimating their performance had already experienced test-score feedback from their past 2 examinations plus they were given bonus points for forming accurate estimations for their last examination. In addition, these students were also given feedback regarding their estimation accuracy. The instructor reported the average over-estimation for the previous tests before the students predicted test 3. In fact, the instructor suggested that a good rule to follow would be to mentally subtract 3.5 from both their predicted and estimated scores (the average overestimation for the first two examinations was 3.675). Apparently, this additional information was helpful because students were now able to predict and estimate their test scores accurately. We expected that once students were able to estimate their test scores accurately, they would continue to do so. This was, however, not the case. Students were accurate in forming their test 3 estimations, but not their test 4 estimations. As expected, the students formed their pre-test predictions for test 4 in respect to their estimations from test 3. What was unexpected was that they failed to lower their

test 4 estimations with respect to the apparent difficulty of the examination (i.e., test 4 produced the lowest mean score of all the examinations, see table 1). What was even more surprising was that while their actual performance was decreasing, their level of estimation, relative to their level of prediction, was increasing. The overall performance from test 3 to test 4 decreased from 43.51 to 39.70, the overall estimation of performance increased from the predicted level of 44.48 to the estimated level of 45.16 (see table 1). We were unable to gather more student feedback regarding test 4 because it was the last examination given in the course. It may have been the case that there was some general misunderstanding regarding test 4, and that this hypothetical misunderstanding produced the estimation failure. We rejected the post-hoc notion because test 4 was similar, in structure, to all the other examinations in the course, plus the fact that there were very few disputes regarding the instructor's grading criteria for any of the previous examinations.

A more intriguing post-hoc hypothesis was that when students were explicitly forewarned about errors due to overestimation they were successful in lowering their test score estimations. When they were not explicitly forewarned they were unsuccessful in lowering their test score estimations.

We were interested in examining the prediction and estimation patterns of students performing above and below the overall test score median. We were able to test for these individual differences with a separate repeated measures (split plot) analysis of variance (ANOVA) for each of the four examinations. A table of predicted, actual and estimated average scores for both students above and below the median are presented in table 3.

TABLE 3
PREDICTED, ACTUAL AND ESTIMATED AVERAGE SCORES FOR STUDENTS ABOVE AND BELOW THE MEDIAN

	RESPONSES	TEST 1	TEST 2	TEST 3	TEST 4
ABOVE	PREDSC	52.93	49.66	48.06	48.53
	ACTSC	49.36	48.53	48.73	45.13
	ESTSC	51.59	50.79	47.40	48.79
BELOW	PREDSC	45.33	43.33	40.33	40.00
	ACTSC	40.56	34.70	38.73	34.13
	ESTSC	43.46	42.00	37.06	41.20

The overall pattern of the results across the four examinations were in line with the results of Moreland, Miller and Laucka (1981), in that students performing above the median were more accurate in estimating their academic performance than students performing below the median. In general both groups tended, to overestimate their

actual test scores with both their pre-test predictions and post-test estimations. However, students above the median were disproportionately closer to the target than students below the median (i.e., the difference between the ESTSC - ACTSC was smaller for good students than for poor students).

For test 1 - conditions of no prior feedback and no incentive for accurate estimation - the overall average for students above and below the median was 51.30 and 43.12 respectively. The difference between these means was very reliable, $F(1,28)=42.66$, $MSe=35.27$, $p<.001$. The average response (across blocks) for the PREDSC, ACTSC and ESTSC was 49.13, 44.96, and 47.53 respectively. This difference was also reliable, $F(2,56)=5.12$, $MSe=25.88$, $p<.01$. Students above and below the median overestimated their test scores at about the same level (the blocks x response interaction was not reliable).

A similar pattern of main effects was obtained for test 2 - conditions of prior test score feedback from test 1 plus an additional incentive was provided for accurate estimation. The overall average for students above the median was 49.66 and the overall average for students below the median was 40.01. The difference between these means was statistically reliable, $F(1,28)=30.02$, $MSe=69.86$, $p<.001$. The main effect for responses (across blocks) was also reliable. Once again students overestimated their actual test scores. The PREDSC was 46.50, the ACTSC was 41.61 and the ESTSC was 46.40, $F(2,56)=13.35$, $MSe=17.5$, $p<.001$. The students above the median were predicting and estimating their test scores with about the same accuracy as for test 1. Students below the median, however, were now proportionally less accurate on test 2 than they were on test 1. This occurred in the face of their feedback from the first examination. Students below the median overestimated their actual performance by only 7% on test 1, but by 21% on test 2. In contrast, students above the median overestimated their actual performance on both tests by only 4%. The blocks x response interaction was very reliable, $F(2,56)=6.26$, $MSe=17.50$, $p<.01$. The mean for test 2 was reliably lower than the mean for test 1 with students below the median contributing disproportionately to the overall drop in the class average. Apparently, these below-the-median students were unable to adjust their predictions to their lower performance. Test 3 produced some unique and interesting results. In this case students were told to mentally subtract a factor for overestimation. More specifically, the instructor told the class to first think of a predicted score and then to subtract 3.5 points from it. They were told to do the same thing with their estimated score after they completed the examination. While, students above the median continued to perform reliably better than students below the median (i.e., good = 48.06, poor = 38.71, $F(1,28)=29.62$, $MSe=66.47$, $p<.001$) they were not any better at predicting or estimating their test

scores. Both students above and below the median estimated their test scores accurately. Both the main effect for responses and the blocks x response interaction were not statistically reliable. Unfortunately, this accurate estimation pattern that emerged for test 3 was not detected following test 4. Test 4, which produced a reliably lower class average than test 3, produced effects similar to test 1. In this case, students received feedback from their past three examinations, and bonus points for accurate estimations. However, they were not given any additional instructions regarding the formation of their predictions and estimations. While good students continued to be reliably superior to poor students (i.e., the overall average for students above the median was 47.48 and for students below the median 38.44, $F(1,28)=48.05$, $MSe=38.30$, $p<.001$), good students and poor students alike overestimated their actual test scores. The overall PREDSC was 44.26, the ACTSC was 39.63 and the ESTSC was 45.00. Differences among these response means were statistically reliable, $F(2,56)=15.34$, $MSe=16.56$, $p<.001$. However, the blocks x response interaction was not reliable. It seems as though changes in the task demands of the examination produced problems for both students above and below the median.

Following the first examination, students were asked to form pre-test predictions and post-test estimations for the average performance of the class. Since students above and below the median tended to overestimate their actual test scores, we wondered whether they would also overestimate the class average. To answer this question we conducted a separate repeated measures analysis of variance (ANOVA) between the dependent measures for tests 2, 3 and 4. The predicted class average (PREDCA), the actual class average (ACTCA, from table 1) and the estimated class average (ESTCA) for tests 2, 3 and 4 are presented in table 4.

TABLE 4
PREDICTIONS AND ESTIMATIONS OF THE AVERAGE PERFORMANCE ON GENERAL
PSYCHOLOGY EXAMINATIONS
SUMMER 1984, N=30

	PREDCA	ACTCA	ESTCA
Test 2	42.96 (72%)	41.66 (69%)	44.57 (74%)
Test 3	42.31 (71%)	43.51 (73%)	41.35 (69%)
Test 4	42.57 (71%)	39.70 (66%)	42.83 (71%)

A separate repeated measures analysis of variance (ANOVA) with students above and below the median as the between-subjects variable, and the pre- and post-test distance from the target (i.e., pre-test distance = $PREDCA - ACTCA$; post-test distance = $ESTCA - ACTCA$), as the within subjects variable was conducted for tests 2, 3 and 4. The results of these analyses indicated that students above and below the median were able to predict the class average with the same degree of precision. There were no reliable differences detected between students above and below the median for any of the three examinations. In fact, the only reliable effect we were able to detect was that for test 2, subjects were closer to the target on the pre-test predictions (i.e., $PREDCA - ACTCA = -1.34$ than their post-test estimations $ESTCA - ACTCA = -2.32$, $F(1,28)=6.71$, $MSe=2.16$, $p<.05$).

Our data, thus far, seem to be consistent with the findings of Moreland, Miller and Laucka (1981) in that students above and below the median form differential assessments of their own test score performance, but are similar in their assessments of the performance of others. While overestimation, as a bias, in test-score prediction and estimation is a pervasive phenomena for both students above and below the median, students below the median are at a greater risk for overestimation bias. It appears that students performing above the median are more cognizant of the task demands involved in test preparation and are therefore less variable in their overall performance than students performing below the median. In other words, good students are able to judge the difficulty of to-be-learned material and adjust their mnemonic strategies during study. Poor students may be less cognizant of the changes in task demands during test preparation and fail to alter their mnemonic strategies during study. Therefore, students performing below the median may be unable to accurately gage the difficulty of a given examination and therefore more variable in their estimation accuracy than students performing above the median.

We believe that while the determinants for overestimation bias are complex and include both cognitive and attributional variables, the attenuation of the effect, however, may be quite simple. We were able to lessen the estimation differences between good and poor students (derived from test 2), by simply forewarning them about their overestimation bias. Apparently, our suggestion to mentally subtract an overestimation factor from their predicted and estimated scores was effective in producing a reliable increase in their estimation accuracy for test 3, but did not generalize to test 4. Maybe if we induced the subtraction rule again, students would have been more accurate in their test 4 estimations. We were encouraged, however, to find that a seemingly simple cognitive manipulation was sufficient to attenuate a fairly robust behavioral phenomena.

Currently, we are attempting to replicate our previous findings with larger classes of general psychology students. We are also trying to experimentally manipulate test score and estimation feedback in order to understand the cognitive parameters involved in overestimation bias. We believe that if students are to become more competent in their cognitive monitoring they must become more accurate and realistic about their academic performance.

REFERENCES

- Biggs, D.A., & Tinsley, D.J. Student-made academic predictions. Journal of Educational Research, 1970, 63, 195-197.
- Edwards, A.L. Experimental Design in Psychological Research. New York: Holt Reinhart, and Winston, Inc., 1972.
- Gadzella, B.M., Cochran, S.W., Parkam, L. & Fournet, G.P. Accuracy and differences among students in their predictions of semester achievement. Journal of Educational Research, 1976, 70, 75-81.
- Keefer, K.E. Characteristics of students who make accurate and inaccurate self-predictions of college achievement. Journal of Educational Research, 1971, 64, 401-404.
- Moreland, R., Miller, J., & Laucka, F. Academic achievement of self-evaluations of academic performance. Journal of Educational Psychology, 1981, Vol. 13, No. 3, 335-344.
- Morrison, T.L., & Morrison, R.L. Personal beliefs, information availability, and judgments of academic success. Psychological Reports, 1979, 45, 483-488.
- Mosteller, F. & Rourke, R.E.F. Study Statistics, Addison-Wesley Publishing Co., 1973.
- Yates, J.W. & Karmos, J.S. Final grade prediction and instructional evaluations at different levels of course achievement. Journal of Instructional Psychology, 1977, 4, 38-40.